PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/68 — A2

(11) International Publication Number: WO 99/05315

(43) International Publication Date: 4 February 1999 (04.02.99)

(21) International Application Number: PCT/GB98/02214

(22) International Filing Date: 24 July 1998 (24.07.98)

(30) Priority Data:

9715942.0	28 July 1997 (28.07.97)	GB
9727103.5	22 December 1997 (22.12.97)	GB

(71) Applicant *(for all designated States except US)*: MEDICAL BIOSYSTEMS LTD. [GB/GB]; The Old Mill, Beaston Cross, Broadhempston, Nr. Totnes, Devon TQ9 6BX (GB).

(72) Inventor; and

(75) Inventor/Applicant *(for US only)*: DENSHAM, Daniel, Henry [GB/GB]; The Old Mill, Beaston Cross, Broadhempston, Nr. Totnes, Devon TQ9 6BX (GB).

(74) Agent: GILL JENNINGS & EVERY; Broadgate House, 7 Eldon Street, London EC2M 7LH (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

(54) Title: NUCLEIC ACID SEQUENCE ANALYSIS

(57) Abstract

The present invention relates to a method for determining the sequence of a polynucleotide, the method comprising the steps of: (i) reacting a target polynucleotide with a polymerase enzyme immobilised on a solid support, and the different nucleotides, under conditions sufficient for the polymerase reaction; and (ii) detecting the incorporation of a specific nucleotide complementary to the target polynucleotide, by measuring radiation.

RE

NUCLEIC ACID SEQUENCE ANALYSIS

Field of the Invention

This invention relates to a method for determining the sequence of a polynucleotide.

Background of the Invention

The ability to determine the sequence of a polynucleotide is of great scientific importance. For example, the Human Genome Project is an ambitious international effort to map and sequence the three billion basis of DNA encoded in the human genome. When complete, the resulting sequence database will be a tool of unparalleled power for biomedical research. The major obstacle to the successful completion of this project concerns the technology used in the sequencing process.

The principal method in general use for large-scale DNA sequencing is the chain termination method. This method was first developed by Sanger and Coulson (Sanger *et al*. Proc. Natl. Acad. Sci. USA 1977; 74: 5463-5467), and relies on the use of dideoxy derivatives of the four nucleoside triphosphates which are incorporated into the nascent polynucleotide chain in a polymerase reaction. Upon incorporation, the dideoxy derivatives terminate the polymerase reaction and the products are then separated by gel electrophoresis and analysed to reveal the position at which the particular dideoxy derivative was incorporated into the chain.

Although this method is widely used and produces reliable results, it is recognised that it is slow, labour-intensive and expensive.

An alternative sequencing method is proposed in EP-A-0471732, which uses spectroscopic means to detect the incorporation of a nucleotide into a nascent polynucleotide strand complementary to a target. The method relies on an immobilised complex of template and primer, which is exposed to a flow containing only one of the different nucleotides. Spectroscopic techniques are then used to measure a time-dependent signal arising from the polymerase

(ii) detecting the incorporation of a specific nucleotide complementary to the target polynucleotide, by measuring radiation.

The radiation may be applied to a sample using a number of techniques, including surface-sensitive detection techniques, where a change in the optical response at a solid optical surface is used to indicate a binding interaction at the surface. In a preferred embodiment of the invention, the technique used is evanescent wave spectroscopy, in particular surface plasmon resonance (SPR) spectroscopy.

In an embodiment of the invention, the nucleotides used in the method include a blocking group at the 3' position, and optionally at the 5' position, which prevents incorporation of the nucleotides into the polynucleotide strand. However, the blocking groups may be selectively removed to allow incorporation to occur. By using the blocked nucleotides, it is possible for the method to be carried out using all the nucleotides present in the reaction at any one time. The selective removal of the blocking groups is carried out in a way that ensures the detection of each incorporated nucleotide. The method may therefore proceed on a "real-time" basis, to achieve a high rate of sequence analysis.

Description of the Drawings

The invention will be described by way of example only with reference to the following drawings, where:

Figure 1 is a schematic illustration of polynucleotide sequence analysis using SPR spectroscopy; and

Figure 2 illustrates the different response signals detected for the polymerisation of each of the different nucleotides.

Figure 3 illustrates the synthesis procedure for the double blocked nucleotides.

Description of the Invention

The present method for sequencing a polynucleotide involves the analysis of the kinetic interaction between a

Publication No. 0648328 (the ntire disclosure of which is incorporated herein by reference).

Nuclear magnetic resonance (NMR) spectroscopy is another preferred method, and measures the magnetic properties of compounds. Nuclei of compounds are energetically orientated by a combination of applied magnetic field and radio-frequency radiation. When the energy exerted on a nucleus equals the energy difference between spin states (the difference between orientation parallel or anti-parallel to the direction of the applied fields), a condition known as resonance is achieved. The absorption and subsequent emission of energy associated with the change from one spin state to the other, is detected by a radio-frequency receiver.

An important aspect of the method of the present invention is the use of a polymerase enzyme immobilised onto a solid support. Immobilisation of the polymerase offers several important advantages for the success of this method. Firstly, the problem of random "noise" associated with measuring energy absorption in soluble molecules is reduced considerably. Secondly, the problem of noise from the interaction of any substrate (e.g. nucleotides) not directly involved with the polymerase is reduced, as the polymerase can be maintained within a specifically defined area relative to the field of measurement. This is particularly relevant if the technique used to measure the changes in radiation requires the measurement of fluorescence, as in TIRF, where background fluorescence increases as the nascent chain grows. Also, if SPR spectroscopy is used, the polymerase reactions are maintained within the evanescent wave field and so accurate measurements can be made irrespective of the size of te polynucleotide. Finally, as neither the target polynucleotide nor the oligonucleotide primer is irreversibly attached to the solid surface, it is relatively simple to regenerate the surface, t allow

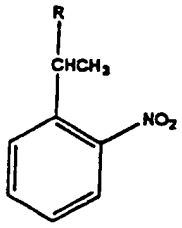
like structure around DNA; and (III) a subassembly of two subunits, τ and γ, used to bind and hydrolyse ATP to form the β-dimer around the DNA.

As a first step in the sequencing process, the target polynucleotide may be brought into contact with an appropriate primer in hybridising/polymerisation buffer. Typically, the buffer will be at a sufficiently high temperature to disrupt (or melt) any secondary structures that exist on the target polynucleotide. On cooling, the primer will anneal to its complement on the target. This sample may then be brought into contact with the immobilised polymerase, to form the target polynucleotide/ polymerase complex.

In one embodiment of the invention, the addition of the nucleotides is controlled so that the different nucleotides are added sequentially to the polymerase/target complex. For example, dGTP may be added and allowed to flow over the polymerase/polynucleotide complex; any incorporation is then detected. Unbound dGTP flows out of the reaction site and a further nucleotide is introduced. In this manner, the detection of a kinetic interaction can be correlated to the particular nucleotide present at that time and the polynucleotide sequence can therefore be determined.

The method may also be carried out with all the different nucleotides present. For this to be carried out successfully, it is necessary for the nucleotides to incorporate a blocking group at least at the 3' position, but preferably at the 3' and 5' positions. The blocking groups may be light-sensitive and can be removed by applying light of a defined wave length, to release the active molecule. If the nucleotides incorporate blocking groups at both the 3' and 5' positions, the blocking groups should be capable of being distinguished on the basis of their spectral absorbancy, i.e. it should be possible to remove selectively one of the blocking groups by applying a specific wavelength of light which does not remove the

A suitable blocking group at the 5' position is 1-(2-nitrophenyl)ethyl group (II):



(II)

wherein R is any suitable functional group, e.g. halogen. This blocking group may be selectively removed at a wavelength of 260 nm.

By way of example, double-blocked nucleotides are injected over the primed target polynucleotide (held in association with a high fidelity polymerase complex), and monochromatic light is focussed upstream of the polymerase at a wavelength sufficient to release the blocking group from the terminal phosphate of each nucleotide. The nucleotides are then able to flow over the bound polymerase, and incorporation into the nascent polynucleotide strand can occur. However, as the blocking group at the 3' position remains bound, only one nucleotide is incorporated. A measurement of the kinetic interaction will therefore provide information as to the particular nucleotide incorporated into the nascent chain. The polymerase used may be a high fidelity polymerase which does not dissociate readily from the target when the reaction stops. Alternatively, a competitive inhibitor may be used to prevent the polymerase dissociating from the target.

After measuring the incorporated nucleotide, a pulse of monochromatic light is focused on the blocking group

concentrated using the ion-filtration technique described by Kirkegaard *et al*, Anal. Biochem. (1972); 50:122.

Immobilisation of the Polymerase

Immobilisation of the polymerase to the sensor chip surface was carried out according to (Jönsson *et al.*, Biotechniques (1991); 11:620-627). Briefly, the sensor chip environment was equilibrated with Hepes buffer (10 mM Hepes, 150 mM NaCl, 0.05% surfactant P20(BIAcore AB, Uppsala, Sweden), pH 7.4). Equal volumes of N-hydroxysuccinimide (0.1 M in water) and N-ethyl-n'-(dimethylaminopropyl) carbodiimide (EDC) (0.1 M in water) were mixed together and injected across the chip (CM5) surface, to activate the carboxymethylated dextran. The polymerase III Subassembly core (160 µl, 500 U) was mixed with 10 mM sodium acetate (100 µl, pH 5) and injected across the activated surface. Finally, residual N-hydroxysuccinimide esters on the sensor chip surface were reacted with ethanolamine (35 µl, 1 M in water, pH 8.5), and non-bound polymerase was washed from the surface. The immobilisation procedure was performed with a continuous flow of Hepes buffer (5 µl/min) at a temperature of 25°C.

Oligonucleotides

Two oligonucleotides were synthesised using standard phosphoramidite chemistry. The oligonucleotide defined as SEQ ID No. 1 was used as the target polynucleotide, and the oligonucleotide defined as SEQ ID No. 2 was used as the primer.

CAAGGAGAGGACGCTGTCTGTCGAAGGTAAGGAACGGACGAGAGAAGGGAGAG SEQ ID No. 1

CTCTCCCTTCTCTCGTC SEQ ID No. 2

The two oligonucleotides were reacted under hybridising conditions to form the target-primer complex.

The primed DNA was then suspended in buffer (20 mM Tris-HCl, pH 7.5, 8 mM $MgCl_2$, 4% (v/v) glycerol, 5 mM

nitrophenyldiazomethane, freshly prepared from 900mg (4 mmol) of 4,5-dimethoxy-2-nitrophenylhydrazone (synthesized by treatment of 6-nitroveraldehyde with hydrazine monohydrate in chloroform by the procedure of Wootton and Trentham, Photochemical Probes in Biochemistry (Nielsen, P.E., Ed,) NATO ASI Ser. C, Vol. 272, p277-296 (1989), was stirred in 15 ml of DMSO at room temperature in the dark for 40 h. Monitoring of the reaction by TLC in a chloroform/methanol (5:1 v/v) solvent system revealed the appearance of a spot with Rf 0.54 corresponding to the caged nucleotide. DMSO, unreacted diazo compound, and reaction products with low polarity were removed by repetitive extraction with 60 ml of ether. The residual material, which, among other substances, contained unreacted nucleotide and the desired product, was dissolved in a minimal amount of chloroform and separated by flash chromatography on a silica column (3 x 30 cm). Elution using 100% chloroform and methanol/chloroform (95:5 v/v) removed the hydrophobic side products of 4,5-dimethoxy-2-nitrophenyldiazomethane from the column. The fractions were dried on a rotary evaporator. 78 mg of the caged product was then lyophilised. The overall yield was 45%. The 3' blocked 4,5-dimethyloxy-2-nitrobenzyl oxycarbonyl dATP was isolated directly with higher purity by preparative reverse-phased HPLC from the crude product.

Stage 2:- Addition of the 5' 1-(2-nitrophenyl)ethyl group to the 3' 4,5-dimethoxy-2-nitrobenzyl oxycarbonyl blocked dATP.

A mixture of 4,5-dimethoxy-2-nitrobenzyl oxycarbonyl 5' dATP (0.4mmol) and approximately 3mmol of 1-(2-nitrophenyl)diazoethane, freshly prepared from 716.7 mg (4mmol) of hydrazone of 2-nitroacetophenone (synthesized by treatment of 2-nitroacetophenone with hydrazine monohydrate in ethanol) and 2.9g (30mmol) of MnO_2 (90%) in 20 ml of chloroform by the procedure of walker *et al* (walker *et al*, Methods Enzymol. 1989; 172:288-301), was stirred in 15 ml of DMSO at room temperature in the dark for 40 h.

the target polynucleotide/ polymerase complex (3) which is held in place by the β-dimer sub-assembly. Since the 3' position on the primer sequence is free to react, polymerisation may take place as a nucleotide is incorporated onto its complement on the target polynucleotide. This incorporation is then detected by the monochromatic p-polarised light of the SPR device. No further polymerisation occurs, since the incorporated nucleotide has a blocking group at the 3' position. Monochromatic light of wavelength 360 nm is then pulsed by the focusing assembly (6) at the site of polymerisation. The high flow rate in the fluidic cell ensures that nucleotides not bound to the polymerase are removed from the cell before sufficient energy has been absorbed to release their 3' blocking groups.

Once the 3' blocking group has been released from the polymerised nucleotide, further polymerisation may occur.

Figure 2 shows the results from the sequencing experiment with each nucleotide incorporated into the nascent chain being detected. The results show a sequence complementary to that of SEQ ID No. 1.

light for a duration different from that required to remove the 3' blocking group.

10. A method according to any preceding claim, wherein step (i) further comprises introducing a competitive inhibitor of the polymerase enzyme.

11. A method according to any preceding claim, wherein the target polynucleotide of step (i) is bound to the polymerase enzyme by a β_2 dimer complex.

12. A method according to any preceding claim, wherein the polymerase is *E. coli* DNA polymerase III or T7 polymerase.

13. A method according to any of claims 1 to 11, wherein the polymerase is Taq polymerase.

14. A method according to any of claims 1 to 11, wherein the polymerase is reverse transcriptase.

15. A method according to any preceding claim, wherein step (ii) comprises detection of a change in resonance signal over time.

16. A method according to any preceding claim, wherein the radiation is electromagnetic.

17. A method according to claim 16, wherein the electromagnetic radiation is in the infra-red spectrum.

18. A method according to any preceding claim, wherein step (ii) comprises using surface plasmon resonance.

19. A method according to claim 16, wherein the electromagnetic radiation is in the radio-frequency spectrum.

20. A method according to claim 19, wherein the incorporation of a nucleotide is detected using NMR.

21. A method according to any preceding claim, wherein the polynucleotide is DNA.

22. A sensor chip comprising a polymerase enzyme immobilised thereon.

23. A nucleotide comprising a blocking group at the 3' position and at the terminal phosphate group of the triphosphate chain, wherein the two blocking groups are removable by monochromatic light of differ nt wavelengths.

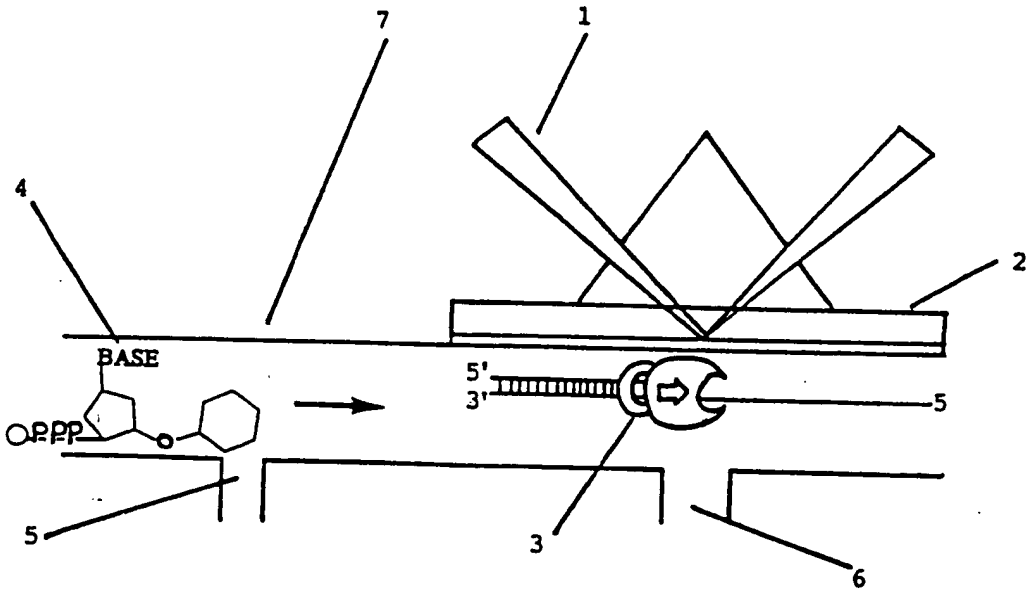


Fig. 1

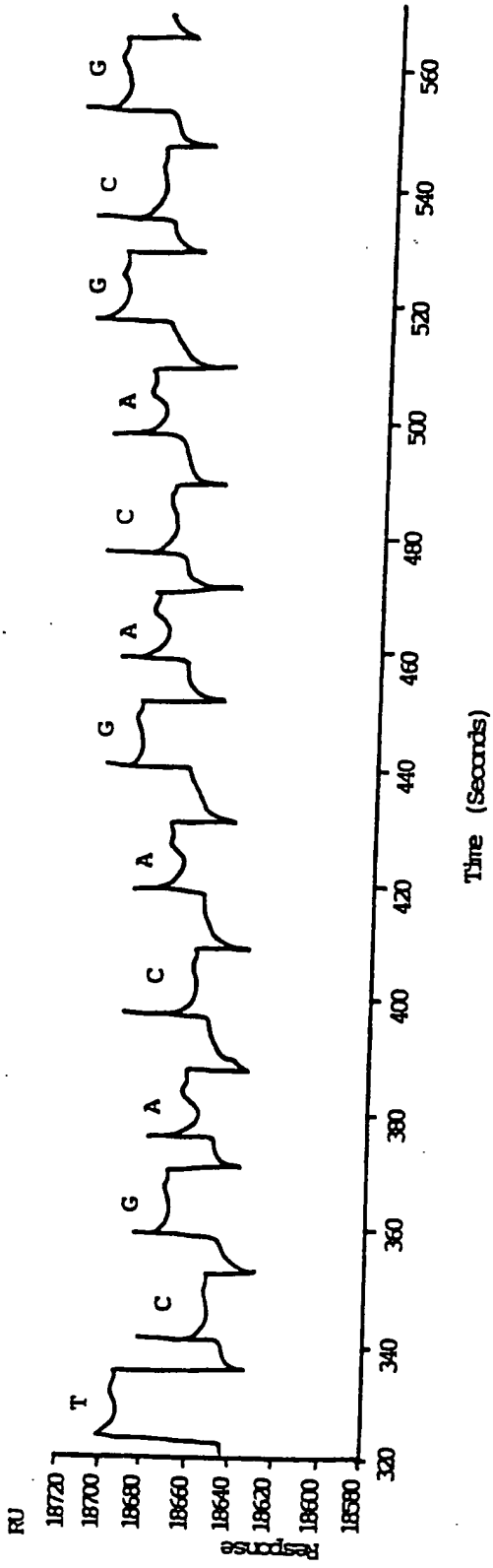

RU
18720
18700
18680
18660
18640
18620
18600
18580
Response
320
340
360
380
400
420
440
460
480
500
520
540
560
Time (Seconds)
T
C
G
A
C
G
A
C
G
A
C
G
C
G

Fig. 2b

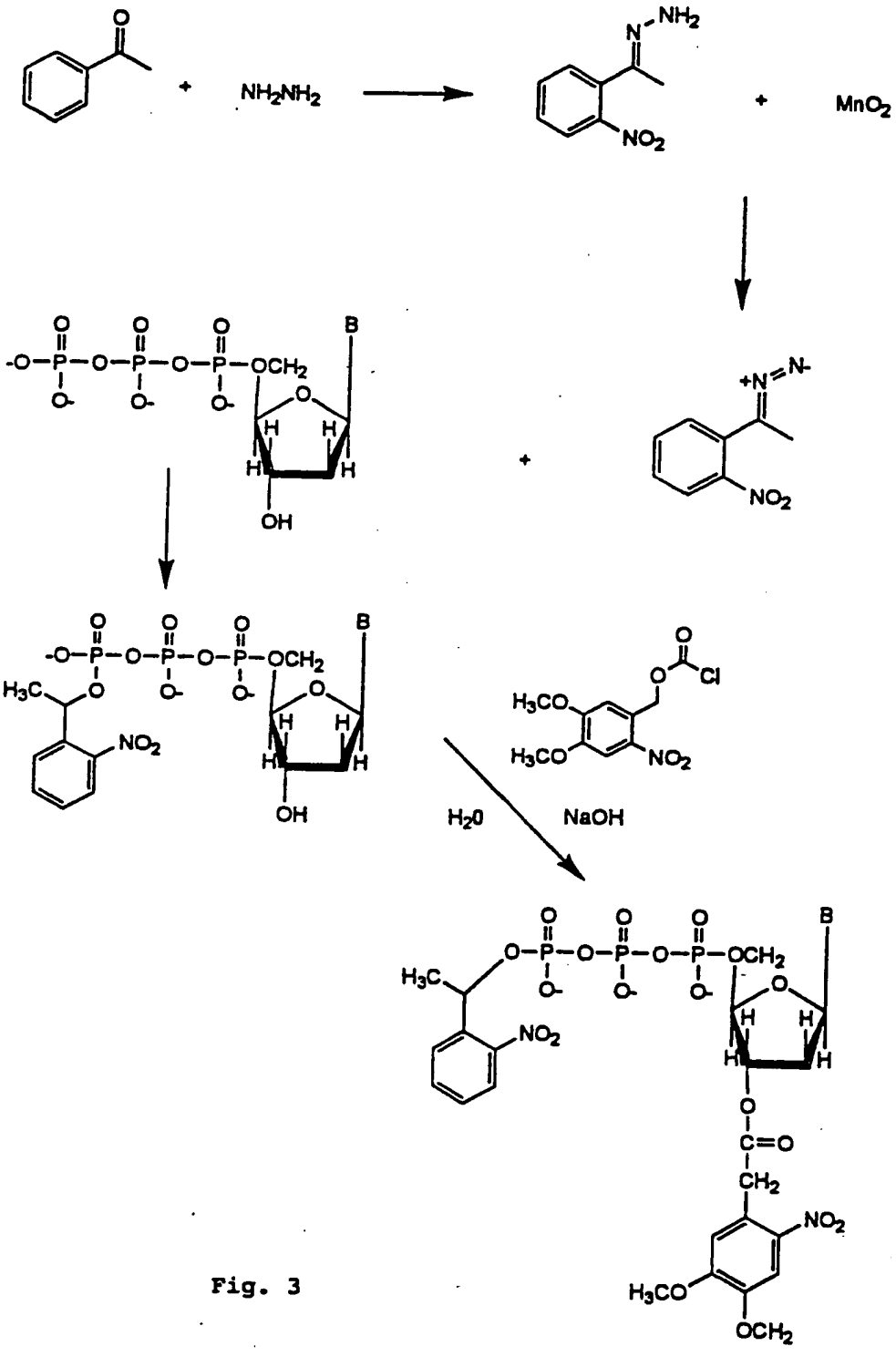


Fig. 3